What is claimed is:

SUB A1

1. A system for pan-genomic determination of three-dimensional macromolecular atomic structures, comprising:

database means for systematically organizing all known structural information into a genomics database of structural information, sequence information and functional information;

bioinformatics means for using said structural information, sequence information and functional information stored in said database to cluster a plurality of known gene products into a plurality of families of homologous sequences;

protein synthesis means for synthesizing for each family, in parallel simultaneously, a plurality of member proteins using genomics information of a plurality of appropriately representative species;

screening means for screening the synthesized proteins to determine ones that are effective;

protein processing means for preparing, purifying and characterizing the screened proteins determined to be effective by said screening means;

crystallization means for crystallizing a plurality of the purified proteins in parallel against a plurality of crystallization screens, and testing a plurality of grown crystals for predetermined diffraction characteristics to determine suitable ones of said plurality of grown crystals;

X-ray crystallography means for performing high-throughput crystallography, said X-ray crystallography means having diffraction measuring means for measuring a suitable crystal for diffraction data, analyzing means for analyzing said diffraction data, means for building an atomic model, and means for refining said model against said diffraction data and storing the refined model in said database;

structure extraction means having means for analyzing the refined model using sequence information of other family members and information of other known three-dimensional structures, means for analyzing for functional motifs and for surface characteristics to define active sites and macromolecular contact sites, and means for defining at least

one class of compounds predicted to have binding potency using the active sites information; and

homology model building means for developing a homology model using an atomic model retrieved from said database,

wherein an ensemble of all known structures is used to further advance an effectiveness of said bioinformatics means.

2. A system according to claim 1, further comprising:

cryoprotection means for freezing said suitable ones of said plurality of crystals which are supplied by said crystallization means,

wherein said suitable crystal is frozen by said cryoprotection means before being measured for said diffraction data by said diffraction measuring means.

3. A system according to claim 1, wherein

said protein synthesis means includes cloning means for cloning for said each family, in parallel simultaneously, cDNAs of said plurality of appropriately representative species into a plurality of expression vectors for a plurality of expressions systems,

said screening means screens for expression constructs obtained from said cloning means to determine ones that are effective, and

said protein processing means processes the expressed proteins determined to be effective by said screening means.

4. A system according to claim 1, wherein

said X-ray crystallography means includes a synchrotron storage ring having undulator beamlines for high-throughput crystallography by a multiwavelength anomalous diffraction method, and

said analyzing means analyzes said diffraction data by a multiwavelength anomalous diffraction phasing method.

5. A system according to claim 4, wherein selenomethionine is incorporated in the synthesized proteins by said protein

synthesis means, and said analyzing means using said multiwavelength anomalous diffraction phasing method analyzes said diffraction data of selenomethionyl proteins.

6. A system according to claim 1, wherein said homology model is used in at least one of target selection, drug design, and design of more appropriate constructs for experimental analysis.

7. A process for pan-genomic determination of three-dimensional macromolecular atomic structures, comprising the steps of:

systematically organizing all known structural information, sequence information and functional information into a genomics database;

clustering a plurality of known gene products into a plurality of families of homologous sequences using at least one bioinformatics tools;

synthesizing for each family, in parallel simultaneously, a plurality of member proteins using genomics information of a plurality of appropriately representative species;

screening the synthesized proteins to determine ones that are effective;

preparing, purifying and characterizing the screened proteins determined to be effective in said step of screening;

crystallizing a plurality of purified proteins in parallel against a plurality of crystallization screens;

testing a plurality of grown crystals for predetermined diffraction characteristics to determine suitable ones of said plurality of grown crystals;

performing high-throughput crystallography, including measuring a suitable crystal for diffraction data, build an atomic model according to an analysis of said diffraction data, refining said model against said diffraction data, and storing the refined model in said database;

analyzing the refined model using sequence information from other family members and information of other known

three-dimensional structures, analyzing for functional motifs and for surface characteristics to define active sites and macromolecular contact sites, and defining at least one class of compounds predicted to have binding potency using the active sites information; and

developing a homology model using computational tools for homology model building and an atomic model retrieved from said database,

wherein said database is updated by using said bioinformatics tools along with said all known structural information and an ensemble of all known structures is used to further advance an effectiveness of said bioinformatics tools.

8. A process according to claim 7, further comprising the step of:

freezing selected ones of said plurality of grown crystals determined to be suitable in said step of testing,

wherein said suitable crystal is frozen before being measured for said diffraction data in said step of measuring.

9. A process according to claim 7, wherein

said step of synthesis includes the step of cloning for said each family, in parallel simultaneously, cDNAs from said plurality of appropriately representative species into a plurality of expression vectors for a plurality of expressions systems,

constructs obtained in said step of cloning are screened for expression to determine ones that are effective, and

the expressed proteins determined to be effective are processed in said step of preparing, purifying and characterizing.

10. A process according to claim 7, wherein

said high-throughput crystallography is performed using a synchrotron storage ring having undulator beamlines along with a multiwavelength anomalous diffraction method, and

said diffraction data is analyzed using a multiwavelength anomalous diffraction phasing method.

11. A process according to claim 10, wherein selenomethionine is incorporated in the synthesized proteins, and said multiwavelength anomalous diffraction phasing method is used to analyze said diffraction data of selenomethionyl proteins.

12. A process according to claim 7, further comprising the step of

using said homology model in at least one of target selection, drug design, and design of more appropriate constructs for experimental analysis.